[Year]

BIOSTATISTICS SBAQ FOR SURGEONS

Biostatistics SBAQ for Surgeons

PREFACE

This book is the result of a personal journey—born out of both frustration and fascination with biostatistics during surgical training. Like many clinicians, I initially found statistical concepts elusive and disconnected from day-to-day practice. Over time, I realised that mastering biostatistics is not just an academic exercise, but a critical skill for interpreting evidence and applying it meaningfully in surgery. This collection of multiple-choice questions reflects that journey. It blends questions of varying difficulty, from foundational principles to more challenging applications, drawn from multiple surgical and clinical specialties. Some questions are repeated in different contexts to reinforce key concepts and ensure high-yield learning.

Whether you're preparing for exams or aiming to sharpen your critical appraisal skills, I hope this book serves as both a practical resource and a reflection of the value of persistence in mastering difficult subjects.

MBBS, MRCSEd
ENT Specialist registrar
Leicester Royal Infirmary

Biostatistics SBAQ for Surgeons

Section 1

1. **To assess whether the likelihood of passing the FRCS exam is greater in National training programme applicants than international applicants.**

   a. Chi-squared (x^2) test
   b. Fisher's exact test
   c. Friedman test
   d. Kruskal-Wallis test
   e. Mann-Whitney U test

2. **Correlation of the blood pressure and age between a group of men and women.**
   a. Pearsons correlation coefficient
   b. Spearman's rank correlation coefficient
   c. Two-sample (unpaired) Student's t-test
   d. Two-way analysis of variance (ANOVA)
   e. Wilcoxon matched pairs test

3. **To determine whether air conduction thresholds change with increasing age.**
   a. Wilcoxon matched pairs test
   b. Spearman's rank correlation coefficient
   c. Two-sample (unpaired) Student's t-test
   d. Pearsons correlation coefficient
   e. Two-way analysis of variance (ANOVA)

4. **To assess whether serum calcium changes twelve hours, one day and three days following total thyroidectomy, is affected by the rank of surgeons.**
   a. Pearsons correlation coefficient
   b. Spearman's rank correlation coefficient
   c. Two-sample (unpaired) Student's t-test
   d. Two-way analysis of variance (ANOVA)
   e. Wilcoxon matched pairs test

5. **To compare the systolic BP measurements of four independent groups (n=80)/group of subjects.**
   a. One-way analysis of variance (ANOVA)
   b. Two-sample (unpaired) Student's t-test
   c. Pearsons correlation coefficient
   d. Two-way analysis of variance (ANOVA)
   e. Wilcoxon matched pairs test

6. **Comparing FRCS-ORL HNS exam scores between the north and south of England.**
   a. Chi-squared (x^2) test
   b. Fisher's exact test
   c. Friedman test
   d. Kruskal-Wallis test
   e. Mann-Whitney U test

Biostatistics SBAQ for Surgeons

7. **To assess the linear analogue scale responses of a group of subjects (n=15) asked to report their perceived pain on 2 consecutive days.**
   a. Pearsons correlation coefficient
   b. One-way analysis of variance (ANOVA)
   c. Two-sample (unpaired) Student's t-test
   d. Two-way analysis of variance (ANOVA)
   e. Wilcoxon matched pairs test

8. **To detect any differences between four consultants' assessments of six trainee candidates performing Microlaryngoscopy**

   a. Kruskal-Wallis test
   b. Fisher's exact test
   c. Chi-squared (x^2) test
   d. Friedman test
   e. Mann-Whitney U test

9. **In a study of pain following tonsillectomy, patients were randomly allocated to receive either an analgesia, or placebo 2 hours pre-operatively. All patients were asked to rate their pain 2 hours after surgery using the scale: 0=no pain, 1=mild, 2=moderate, 3=severe. Analyse the results.**

   a. Fisher's exact test
   b. One-way analysis of variance (ANOVA)
   c. Two-sample (unpaired) Student's t-test
   d. Mann-Whitney U test
   e. Wilcoxon matched pairs test

10. **The number of GP consultation in the first 10 year of life was obtained for a group of 200 children. Their parents were asked to complete a parenting skills questionnaire from which a skills score (an integer value ranging from 0-25) was determined. Which one of the above statistical test would be the most appropriate for analysis to assess whether low parenting skills are related to a higher number of GP consultations?**

    a. Pearsons correlation coefficient
    b. Spearman's rank correlation coefficient
    c. Two-sample (unpaired) Student's t-test
    d. Two-way analysis of variance (ANOVA)
    e. Wilcoxon matched pairs test

11. **You are critically appraising a paper looking at the effect of tranexamic acid on patient outcome following functional endoscopic sinus surgery. The results are presented with p-values. At what p-value is the result statistically significant?**
    a. <1
    b. <0.9
    c. <0.5
    d. <0.1
    e. <0.05

Biostatistics SBAQ for Surgeons

12. **You are reviewing a paper that is assessing a novel test for the diagnosis of Meniere's Disease. The paper presents the sensitivity and specificity of the test. Which statement is true regarding these statistical measures?**
    a. Sensitivity increases with the increasing prevalence of a condition
    b. The sensitivity is calculated as the number of true negatives divided by all of those without the disease
    c. Specificity decreases with decreasing prevalence of a condition
    d. Sensitivity measures the proportion of patients who have the disease who are identified correctly by the test
    e. The specificity is calculated as the number of true positives divided by all of those with the disease

13. **A placebo-controlled randomised trial of antibiotics in acute sinusitis is reported in JLO. The patients who received antibiotics were reported to have a 77% cure rate compared to 74% in the placebo group. What is the ARR?**
    a. 3%
    b. 5%
    c. 33%
    d. 0.3%
    e. 26%

14. **A placebo-controlled randomised trial of antibiotics in acute sinusitis is reported in JLO. The patients who received antibiotics were reported to have a 77% cure rate compared to 74% in the placebo group. What is the NNT?**
    a. 77
    b. 74
    c. 66
    d. 33
    e. 3

15. **Comparing average blood pressures between a group of men and women with epistaxis.**
    a. Chi-squared (x^2) test
    b. 1-way ANOVA
    c. Paired T-test
    d. Unpaired T-test
    e. Mann-Whitney U test

16. **Medical statistics deals with the analysis, interpretation and presentation of data. Being able to compare datasets for statistical differences is key for evidence-based healthcare. Which of the following statistical tests are appropriate for analysing parametric (normally distributed) data?**

    a. Chi-squared
    b. Kruskall-Wallis
    c. Mann-Whitney
    d. Friedman's two-way ANOVA
    e. Student's t-test

Biostatistics SBAQ for Surgeons

17. **The way datasets are distributed has a significant impact on the statistical tests that are required. The Gaussian distribution is used in statistical analysis to determine two limits within which an observation will fall. Which of the following options is correct about Gaussian distribution?**

    a. The mean value is an accurate way of assessing the central value of a data set, even when the data are skewed

    b. The standard deviation (SD) is a measure of how accurately the calculated mean approaches the true population mean

    c. The ANOVA test may be used to compare this set of values with one further set provided they are also normally distributed

    d. The modal value is always equal to the mean

    e. 25% of the values will have numerical values that are smaller than the mean value minus 5.96 standard deviations

18. **You are at an interview for an ST3+ appointment and are being asked a question on the different levels of evidence that you know of. Which of the following is an example of level 2 evidence-based medicine?**
    a. Expert opinion
    b. Case series
    c. Cohort study with a high risk of confounding or bias
    d. Systematic review of randomised controlled trials
    e. Randomised controlled trial (RCT) with a low risk of bias

19. **With regards to measures of central tendency in statistics, which of the following statements is true?**
    a. The mean is always a good measure of central tendency
    b. The mode is the sum of all the variables divided by the number of variables
    c. There is only one mode in any given distribution
    d. The median is not the middle of the distribution
    e. The mean, mode and median are equal in normal distributions

20. **The following statements on clinical trials and evidence- based medicine is correct?**
    a. Treatments supported by a single RCT are said to have level 1a evidence

    b. An intention to treat analysis means that patients outcomes are analysed according to the treatment that they received, as opposed to the group to which they were randomised

    c. A study has a type II error if it finds an effect or difference when none exists

    d. An absolute risk reduction (ARR) is the risk in the control group minus the risk in the treated group, and is the inverse of the number needed to treat (NNT)

    e. A p-value of less that 0.05 means that the probability of a result occurring by chance is less than 1 in 5.

Biostatistics SBAQ for Surgeons

21. An ENT Surgeon wants to determine whether a history of childhood tonsillectomy affects rates of adult atopic disease. He assembles a group of adults with and without atopic disease and then compares the rates of tonsillectomy by reviewing their records. This is an example of which type of study?

a. Prospective observational study
b. Retrospective observational study
c. Case-control study
d. Case-series study
e. Randomised controlled study

22. Visual analogue scale responses of a group of subjects (n = 16) who were asked to report their perceived contentment on 2 consecutive days. Which is the most suitable statistical test?

a. Wilcoxon matched pairs test
b. Mann-Whitney U test
c. Chi-Squared (X2) test
d. One Sample (paired) Student's t-test
e. Two-sample (unpaired) Student's t-test

23. With regards study validation, which one of the following statements is true?

a. Specificity indicates how good the test is at picking up people who have the condition
b. Sensitivity indicates how good the test is at correctly excluding people without the condition
c. The negative predictive value (NPV) is the pre test probability that if a person's test is negative, that he or she has the disease
d. Positive predictive value (PPV) is the post-test probability of a positive result and is the probability of someone having the disease if they have a positive test result
e. The prevalence of the condition is the number with disease in the sample size of the population size

24. In an observational cohort study of treatment of sinusitis, patients are treated with either antibiotics or surgery. Failure to recognize that the patients who receive surgical intervention generally have more severe symptoms represents a problem with:

a. Selection bias
b. Confounding bias
c. Detection bias
d. Comorbidity bias
e. Intervention bias

25. The variation of weight over a period of 3 months in 600 obese subjects on a calorie-controlled weight-reducing diet. Which is the most suitable statistical test?

a. Kruskal-Wallis test
b. Mann-Whitney U test
c. McNemar
d. One Sample (paired) Student's t-test
e. One-way analysis of variance (ANOVA)

26. A clinician reviews the medical records of all patients who had tonsillectomy over the past 10 years and records the frequency of primary haemorrhage. Follow-up is available for all subjects. The haemorrhage rate is reported using

a. prevalence, because the method of data collection is retrospective
b. incidence, because the method of collection is prospective
c. prevalence because the direction of inquiry is retrospective
d. incidence because the direction of inquiry is prospective
e. Survival analysis because some observations are censored.

Biostatistics SBAQ for Surgeons

27. **Evidence obtained from one randomised controlled trial provides what level of evidence?**
    a. Level Ib
    b. Level Ia
    c. Level IIa
    d. Level IIb
    e. Level III

28. **A new screening for a lethal disease is going to be introduced in the whole population. The prevalence of the disease is 0.5 percent. The screening process has a risk of carcinogenesis of 1 in 4 million. Which statement do you consider most important?**

    a. The sensitivity of the screening test is 97%
    b. The specificity of the screening test is 97%
    c. The treatment of early detected disease improves the cure rate in 75% of the patients
    d. The cost of each screening test is 35p
    e. The condition affects only a population of over 75 years old.

29. **When the change in hearing levels is assessed after ossicular reconstruction, a 95 percent confidence interval aids in data interpretation because it**
    a. Gives the range of results consistent with the data
    b. Estimates the accuracy of observed results
    c. Adjusts for systematic error that may have occurred when results were measured
    d. Defines the variability of observed data relative to the mean
    e. Ensures adequate statistical power

30. **The most important aspect of analysing and interpreting medical data is?**

    a. Choosing the right statistical test for the right dataset
    b. Recognising uncertainty and quantifying error rates
    c. Avoiding the multiple p-value problem by using multivariate techniques
    d. Reporting statistical power whenever a significant p-value is obtained
    e. Reporting statistical power whenever a non- significant p-value is obtained

31. **In designing a trial comparing primary surgery against primary chemoradiotherapy for head and neck cancer, which of the following statements concerning randomisation is true?**
    a. The number of patients in each arm will be the same
    b. The clinician can favour one treatment over another if the patient expresses choice
    c. The clinician will not be aware of which treatment the patient has received
    d. The patients in each arm should have similar prognostic factors
    e. Randomisation ensures that there is no bias between the groups

32. **A new screening test for a squamous cell carcinoma is developed. In total, 150 people were tested. Within the whole study, 50 patients were known to have the disease. When testing was started, 40 patients who tested positive had the disease, while ten patients who tested positive did not have the disease. Which of the following statements is correct?**
    a. The specificity of the test is 80%
    b. The sensitivity of the test is 80%
    c. The positive predictive value of the test is 90%
    d. The negative predictive value of the test is 80%
    e. The prevalence of squamous cell carcinoma in this study is 40%

Biostatistics SBAQ for Surgeons

33. **Concerning the p-value, which of the following statements is true?**
    a. p-value of 0.001 with a significance level set at 0.05 indicates that the null hypothesis is wrong
    b. For a p-value to be significant it must be a value less than 0.05
    c. A type l error is to reject the null hypothesis when it is true
    d. A p-value of 0.0001 indicates a highly significant clinical finding
    e. E. The null hypothesis should be rejected if the p-value is 0,1 when the significance level has been set at less than 0.05

34. **Concerning statistical testing, which of the following statements is false?**
    a. Parametric tests make an assumption about the population distribution
    b. Standard error of a population is calculated as the standard deviation divided by the square root of the population size
    c. Specificity relates to the number of true positives in a test
    d. A type 1 error is to reject the null hypothesis when it is true
    e. A type 2 error is to accept the null hypothesis when it is false

35. **Concerning anticancer clinical drug trials, which of the following statements is correct?**
    a. Phase 1 clinical trials usually involve large numbers of patients to test for anticancer properties
    b. Phase 2 trials will aim to ascertain the response of the cancer to the treatment
    c. A Phase 3 trial testing a new cancer drug will assess its cost-effectiveness
    d. All cancer patients should be included in clinical trials
    e. Patient consent is not required for the inclusion of patients with advanced cancer into phase 1 trials only

36. **A researcher reports that a new antibiotic is more effective than an established standard for bacterial sinusitis, p = 0.015. Which of the following statement is true? (ENT TZAR)**

    a. The new antibiotic increased relative efficacy by 15%
    b. The new antibiotic increased absolute efficacy by 15%
    c. The likelihood of a type I error is 1.5%
    d. The likelihood of a type Il error is 1.5%
    e. The likelihood of adequate statistical power is 98.5%

37. **A randomised controlled trial is conducted to evaluate a novel therapy. What level of evidence does this represent?**
    a. 1a
    b. 1b
    c. 2a
    d. 2b
    e. 3

38. **A Rhinology SpR is doing a research looking at the efficacy of a new topical nasal steroid comparing it with placebo. Each group has 20 participants. Acoustic rhinometry will be used to assess nasal patency, with mean values and standard deviation measured. When assessing the Null Hypothesis, which of the following tests would be most appropriate?**
    a. Analysis of variance
    b. Chi-Square test
    c. Fisher's exact test
    d. Spearman Rank Correlation
    e. Student t-test

Biostatistics SBAQ for Surgeons

39. **Parametric statistical tests can be used for the following?**
    a. Categorical data with non-normal distribution
    b. Categorical data with normal distribution
    c. Categorical data with bimodal distribution
    d. Continuous data with normal distribution
    e. Continuous data with non-normal distribution

40. **In an observational cohort study of treatment of sinusitis, patients are treated with either antibiotics or surgery. Failure to recognize that the patients who receive surgical intervention generally have more severe symptoms represents a problem with:**
    a. Selection bias
    b. Confounding bias
    c. Detection bias
    d. Comorbidity bias
    e. Intervention bias

41. **A clinician reviews the medical records of all patients who had tonsillectomy over the past 10 years and records the frequency of primary haemorrhage. Follow-up is available for all subjects. The haemorrhage rate is reported using**
    a. prevalence, because the method of data collection is retrospective
    b. incidence, because the method of collection is prospective
    c. prevalence because the direction of inquiry is retrospective
    d. incidence because the direction of inquiry is prospective
    e. Survival analysis because some observations are censored

42. **Which of following study designs is most appropriate to investigate if exposure to environmental tobacco smoke is associated with the development of esthesioneuroblastoma?**

    a. Cross-sectional study
    b. Randomized controlled trial
    c. Case control study
    d. Cohort study
    e. Audit

43. **An ENT surgeon reports that a new antibiotic is more effective than an established one for bacterial acute otitis media, p = 0.015.**
    a. The probability that the ENT surgeon is incorrect is 1.5%
    b. The new antibiotic increased absolute efficacy by 15%
    c. The likelihood of a type I error is 1.5%
    d. The likelihood of a type II error is 1.5%
    e. The likelihood of adequate statistical power is 98.5%

44. **This test is used to compare the medians of two groups of independent non-parametric numerical data**
    a. Spearman correlation
    b. Mann–Whitney U test
    c. Chi-square
    d. T-test
    e. Pearson correlation

Biostatistics SBAQ for Surgeons

45. Which of the following statements is INCORRECT when referring to a p-value?

a. Assesses the strength of evidence against the null hypothesis.
b. Is the probability that the null hypothesis is true.
c. A p-value less than 0.05 means that the 95% confidence interval does not contain the null value.
d. The p-value does not indicate the size or importance of the observed effect.
e. When interpreting the results of an analysis, the p-value should always be considered alongside the confidence interval.

46. Publication bias is best detected in a meta-analysis using which of the following graphical forms?

a. Funnel plot
b. Kaplan–Meier curve
c. Receiver operator characteristic curve
d. Forest plot
e. Histogram

47. A controlled study, which sequentially allocates consecutive patients to either bipolar tonsillectomy or coblation tonsillectomy to compare short-term outcomes between the two treatment modalities, provides which level of evidence?

a. Level 1
b. Level 2
c. Level 3
d. Level 4
e. Level 5

48. In a non-randomised controlled study evaluating the hypothesis that endoscopic dacryocystorhinostomy, as compared with external dacryocystorhinostomy, reduces the length of hospital stay; the null hypothesis was inappropriately accepted. This is known as?

a. Type I error
b. False positive error
c. Power of a study
d. Publication bias
e. Type II error

49. Which one of the following study designs would be best suited to assess the incidence of secondary lung malignancy in patients receiving external beam radiotherapy for laryngeal cancer?

a. Observational study
b. Randomised controlled study
c. Case-control study
d. Cohort study
e. Cross-sectional study

50. Which of the following statements regarding statistical data is true?

a. The mean is a measure of data spread.
b. The mode is a measure of data location.
c. The median is the most commonly occurring bauelin a data set.
d. A median can be used with nominal data.
e. A mean cannot be used with odinal data.

Biostatistics SBAQ for Surgeons

51. **In a study evaluating side effects of medical therapy for allergic rhinitis, men are randomly allocated to pseudoephedrine (n = 100) or placebo (n = 100) and subsequently monitored for progression to an episode of acute urinary retention (AUR). In the placebo arm 25 men developed AUR compared with 20 in the pseudoephedrine arm. What is the probability of progressing to AUR in those taking the pseudoephedrine?**
    a. 0.2
    b. 5%
    c. 1/4
    d. 0.5
    e. 25%
52. **In the same study as described above, what are the odds of developing AUR in the placebo arm?**
    a. 0.30
    b. 0.31
    c. 0.32
    d. 0.33
    e. 0.34

53. **Which of the following statements regarding statistical data is true?**

    a. Sex and ethnicity are examples of ordinal data.
    b. The mode is the middle value in a data set.
    c. Weight and height are discrete variables.
    d. Standard deviation is a measure of data spread.
    e. GCS score and age are examples of nominal data

54. **A study into the effectiveness of a novel intratympanic drug for the treatment of sudden sensorineural hearing loss reports that the difference in hearing outcomes between the novel drug-treated group and the standard-of-care treatment group was statistically significant (p-value = 0.002). Which of the following statements is correct?**
    a. A p-value <0.05 means the difference is clinically meaningful
    b. The higher the p-value the more clinically meaningful the result
    c. The threshold for statistical significance is always p<0.05
    d. A p-value of 0.002 means that the result has a 99.8% chance of being true
    e. A p-value of 0.002 means the difference between the groups would have a 0.2% chance of occurring by chance

55. **A research study compared the olfactory dysfunction in men and women with COVID-19. At 6 months, the odds ratio of olfactory dysfunction in women compared to men was 3.1 (95% CI 1.2 to 9.4, p-value = 0.01). Which statement is correct?**
    a. Men experienced more olfactory dysfunction than women 6 months after COVID-19 infection
    b. Women experienced approximately three times the severity of olfactory dysfunction compared to men 6 months after COVID-19 infection
    c. The odds ratio of olfactory dysfunction in women is calculated by dividing the number of women with olfactory dysfunction by the number of men with olfactory dysfunction
    d. The odds ratio indicates that more women than men experienced olfactory dysfunction 6 months after COVID-19 infection
    e. A lower p-value means a more clinically significant difference in olfactory dysfunction

Biostatistics SBAQ for Surgeons

56. You are wanting to investigate whether the choice of intranasal steroid affects the symptom control of chronic rhinosinusitis. You decide this is best approached by a randomised controlled trial (RCT). Which of the following statements is correct?

a. RCTs are better suited for retrospective analysis
b. RCTs are particularly suited for the evaluation of rare outcomes such as the development of glaucoma with intranasal steroid usage
c. Randomisation by online randomisation software can minimise the effect of confounding factors
d. Using the first letter of the patient's name as a method of randomizing patients would remove the chance of allocation bias
e. An RCT design for this study is appropriate since RCTs require shorter follow-up periods

57. In a meta-analysis conducted to evaluate the efficacy of a tonsillectomy for treating recurrent tonsillitis, the researchers include studies with varying surgical techniques. Which statistical method is commonly used in meta-analysis to address this heterogeneity and provide a pooled estimate of treatment effect?

a. Sensitivity analysis
b. Funnel plot analysis
c. Fixed-effects model
d. Random-effects model
e. Stratified analysis

58. You are explaining the use of vestibular evoked myogenic potentials (VEMPs) in the investigation of superior semicircular canal dehiscence (SSCD) to a medical student. You explain that the test at your centre has a sensitivity of 85%. The medical student asked what this means. What is another way to explain this?

a. 85% of those with a positive test have SSCD
b. The test correctly identifies 85% of patients without SSCD
c. The test has a 15% probability to return a false positive result
d. 85% of those with SSCD will have a positive test result
e. The test is not reliable for diagnosing SSCD

59. You are designing a randomised controlled trial to assess the apnoea-hypopnea index (AHI) after cold steel and coblation tonsillectomy in children diagnosed with obstructive sleep apnoea.A sample size of 100 was calculated based on an 80% power to detect a 3-point difference in AHI. You were able to recruit 90 patients to the trial and you report that AHI was significantly lower in the coblation group (p-value = 0.04). Which of these statements is correct?

a. It is likely that there is a clinically meaningful difference in the AHI between the groups
b. The power to detect a difference of 3 points was lower than 80%
c. The trial was adequately powered to reject the null hypothesis
d. The difference in AHI between the 2 groups is significantly larger than 3 points
e. The trial should have been stopped if the number of recruited patients was less than 100

60. In a population-based study investigating the prevalence of laryngeal cancer among individuals who smoke and those who do not, what type of epidemiological study design is most suitable for assessing the association between smoking and laryngeal cancer risk?

a. Case-control study
b. Cross-sectional study
c. Cohort study
d. Randomised controlled trial
e. Ecological study

Biostatistics SBAQ for Surgeons

61. For which of the following situations would an independent sample t-test be appropriate?

a. Comparing patient body weight between those with and without tonsillitis
b. Comparing pain on a Likert scale before and after tonsillectomy
c. Comparing patient satisfaction before and after tonsillectomy for recurrent tonsillitis
d. Comparing the proportion of patients with recurrent sore throats in those who have had or have not had tonsillectomy
e. Comparing patient systolic blood pressure before and after tonsillectomy

62. You are analysing the results of your personal series of thyroidectomies. You find that the Pearson correlation between patient satisfaction and length of stay is -0.7 (p = 0.01). Which of the following is correct?

a. Longer patient stays were associated with greater patient satisfaction
b. A correlation coefficient of -0.7 indicates a strong statistical significance
c. There is a strong negative correlation between the length of stay and patient satisfaction
d. There is a weak correlation between the length of stay and patient satisfaction
e. A p-value of 0.05 would suggest a more statistically significant result than a p-value of 0.01

63. A journal article reports the result of a randomised controlled trial comparing vocal cord injection with a new biomaterial to a placebo. They report the difference in voice outcome between the intervention and control group as measured on a 0–10 cm visual analogue scale (VAS) was 1.2 cm (95%CI -1.2 cm to 3.6 cm). Which of these statements is correct?

a. 95% of the group treated with the new material had voice improvement of between -1.2 cm and 3.6 cm
b. 5% of the control group had voice improvement of between -1.2 cm and 3.6 cm
c. The improvement in voice outcomes of the new material could have been due to chance
d. A VAS is an inappropriate tool for use with this type of data
e. Estimating 99% confidence intervals would give a more accurate estimate

64.

Biostatistics SBAQ for Surgeons

Part 2

The following 2x2 table shows the results of validation study of urine glucose testing for diabetes against the gold standard of glucose tolerance test.

		Results of Glucose tolerance test	
		Diabetes Positive	Diabetes Negative
		27 Subjects	**923 subjects**
Results of urine test for Glucose	Glucose Present	True Positive	False positive
	13 subjects	**6**	**7**
	987 subjects	**21**	**966**
	Glucose Absent	False negative	True Negative

65. What is the specificity of the urine test for diabetes?

a. 22.2%
b. 99.3%
c. 46.2%
d. 97.8%
e. 97.2%
f. 32
g. 0.78

66. What is the sensitivity of the urine test for diabetes?

a. 22.2%
b. 99.3%
c. 46.2%
d. 97.8%
e. 97.2%
f. 32
g. 0.78

67. What is the positive predictive value of the urine test for diabetes?

a. 22.2%
b. 99.3%
c. 46.2%
d. 97.8%
e. 97.2%
f. 32
g. 0.78

68. What it the likelihood ratio of a negative test for diabetes?

a. 22.2%
b. 99.3%
c. 46.2%
d. 97.8%
e. 97.2%
f. 32
g. 0.78

Biostatistics SBAQ for Surgeons

Theme: Types of Studies

69. A prospective comparison of two treatments:
a. Case-control study
b. Cohort study
c. Randomized controlled trial
d. Systematic review
e. Case series study

70. A retrospective study of patients with a disease in an attempt to identify risk factors for that disease

a. Case-control study
b. Cohort study
c. Randomized controlled trial
d. Systematic review
e. Case series study

71. A survey to examine the prevalence of a disease

a. Case-control study
b. Cohort study
c. Randomized controlled trial
d. Systematic review
e. Case series study

72. A prospective study in which subjects, initially disease-free, are followed up

a. Case-control study
b. Cohort study
c. Randomized controlled trial
d. Systematic review
e. Case series study

Theme: Levels of evidence

73. What level of evidence is an Individual randomised controlled trial (RCT)?
a. Level 1a
b. Level 1 b
c. Level 1 c
d. Level 2a
e. Level 2 b

74. What level of evidence is Case series?
a. Level 1
b. Level 2
c. Level 3
d. Level 4
e. Level 5

Biostatistics SBAQ for Surgeons

75. What level of evidence is a Systematic review (with homogeneity) of RCTs?

a. Level 1a
b. Level 1 b
c. Level 1 c
d. Level 2a
e. Level 2 b

76. What level of evidence is Expert opinion?

a. Level 1
b. Level 2
c. Level 3
d. Level 4
e. Level 5

77. Input variable and outcome variable are both ordinal

a. Chi-square test
b. Student's t-test
c. Mann-Whitney U test
d. McNemar's test
e. Spearman rank correlation
f. Logistic regression
g. Kruskal-Wallis test

78. Input variable is nominal and outcome variable is non-normal

a. Chi-square test
b. Student's t-test
c. Mann-Whitney U test
d. McNemar's test
e. Spearman rank correlation
f. Logistic regression
g. Kruskal-Wallis test

79. Input variable is nominal and outcome variable is normal

a. Chi-square test
b. Student's t-test
c. Mann-Whitney U test
d. McNemar's test
e. Spearman rank correlation
f. Logistic regression
g. Kruskal-Wallis test

Biostatistics SBAQ for Surgeons

80. Input variable is nominal and outcome variable is quantitative discrete

a. Chi-square test
b. Student's t-test
c. Mann-Whitney U test
d. McNemar's test
e. Spearman rank correlation
f. Logistic regression
g. Kruskal-Wallis test

81. The sample mean plus or minus 1.96 times its standard error;
a. Standard deviation
b. Standard error of the mean
c. Type I error
d. Type ll error
e. 95% confidence interval
f. Reference range
g. Standard error

82. Standard deviation/√n;
a. Standard deviation
b. Standard error of the mean
c. Type I error
d. Type ll error
e. 95% confidence interval
f. Reference range
g. Standard error

83. An estimate of variability of observations;
a. Standard deviation
b. Standard error of the mean
c. Type I error
d. Type ll error
e. 95% confidence interval
f. Reference range
g. Standard error

84. A measure of precision of an estimate of a population parameter;
a. Standard deviation
b. Standard error of the mean
c. Type I error
d. Type ll error
e. 95% confidence interval
f. Reference range
g. Standard error

Biostatistics SBAQ for Surgeons

Statistical terms

For each of the descriptions below, choose the most appropriate statistical term from the list above.

85. Is calculated as 1 divided by the absolute risk difference;

- **a.** Absolute risk difference
- **b.** Relative risk ratio
- **c.** Number needed to treat
- **d.** Number needed to harm
- **e.** Prevalence
- **f.** Sensitivity
- **g.** Incidence
- **h.** Negative predictive value
- **i.** Specificity

86. Is defined as the number of patients who must be treated in order to prevent one adverse event;

- **a.** Absolute risk difference
- **b.** Relative risk ratio
- **c.** Number needed to treat
- **d.** Number needed to harm
- **e.** Prevalence
- **f.** Sensitivity
- **g.** Incidence
- **h.** Negative predictive value
- **i.** Specificity

87. Cannot be calculated by a cross-sectional study design;

- **a.** Absolute risk difference
- **b.** Relative risk ratio
- **c.** Number needed to treat
- **d.** Number needed to harm
- **e.** Prevalence
- **f.** Sensitivity
- **g.** Incidence
- **h.** Negative predictive value
- **i.** Specificity

Biostatistics SBAQ for Surgeons

88. You develop a new test for Ménière's disease. You discover that the test correctly identifies 86% of those who actually have Ménière's disease. What is this statistic called?

a. Sensitivity
b. Positive predictive value
c. Specificity
d. False negative
e. Odds ratio
f. False positive
g. Relative risk
h. Incidence
i. Negative predictive value
j. Prevalence

89. You investigate your test for Ménière's disease further and find that 74% of those who test negative for Ménière's are truly disease-free. What is this statistic called?

a. Sensitivity
b. Positive predictive value
c. Specificity
d. False negative
e. Odds ratio
f. False positive
g. Relative risk
h. Incidence
i. Negative predictive value
j. Prevalence

90. Your new test for Ménière's identifies 38% of patients who actually have appendicitis as being negative for Ménière's disease. What is this statistic called?

a. Sensitivity
b. Positive predictive value
c. Specificity
d. False negative
e. Odds ratio
f. False positive
g. Relative risk
h. Incidence
i. Negative predictive value
j. Prevalence

91. You are keen to determine whether smoking is a risk factor for a new disease that has been identified. You take the proportion of smokers who develop the disease and divide this by the proportion of non-smokers who develop the disease. What is this statistic called?

a. Sensitivity
b. Positive predictive value
c. Specificity
d. False negative
e. Odds ratio
f. False positive
g. Relative risk
h. Incidence
i. Negative predictive value
j. Prevalence

Biostatistics SBAQ for Surgeons

92. This new disease affects 28 new people per 100,000 per year. What is this statistic called?

a. Sensitivity
b. Positive predictive value
c. Specificity
d. False negative
e. Odds ratio
f. False positive
g. Relative risk
h. Incidence
i. Negative predictive value
j. Prevalence

Biostatistics SBAQ for Surgeons

Surgical Statistics

93. A new blood test has been produced to help diagnose suspected tumours of the membranous urethra. Of a study sample of 100 patients, 20 are known to have cancer of the membranous urethra. The blood test suggested that 30 people had the tumour, but of these only 15 were true positives. What is the positive predictive value of this test?

a. 25%
b. 50%
c. 75%
d. 90%
e. 100%

94. Tented T waves on electrocardiograms are thought to reflect hyperkalemia. A new proposed definition of tented T waves is 'a T wave that is taller than 2 large squares. A study looking at the feasibility of this definition found 1000 ECGs: 200 with tented T waves (according to the proposed definitions) and 800 without. The study found that of the 200 people with ECGs with tented T waves, 100 had hyperkalaemia. Of the 800 people without tented T waves, a further 100 had hyperkalaemia.

What is the specificity of this definition of tented T waves with relation to hyperkalaemia?

a. 12.5%
b. 25.0%
c. 50.0%
d. 87.5%
e. 95.0%

95. A new form of CT scan is being piloted to help detect scaphoid fractures. A study looked at 600 people: 300 people with scaphoid fractures and 300 without. The trial produced 250 positive results and 350 negative results. Of the 250 positive results, there were no false positives, and of the 350 negative results, 50 were false negatives. What is the sensitivity of CT scanning in detecting scaphoid fractures in this study?

a. 17%
b. 33%
c. 50%
d. 83%
e. 100%

Biostatistics SBAQ for Surgeons

96. A new blood test is being developed to help diagnose sarcoidosis. In a trial 100 patients have been tested. The trial produces 20 positive results and 80 negative results. Of the 20 positive results, 10 of them are false positives. Of the 80 negative results, 10 of them are false negatives. What is the sensitivity of this test?

a. 10%
b. 25%
c. 50%
d. 75%
e. 90%

97. A new tumour marker is being tested for use in cholangiocarcinomas. In a trial 40 patients have been tested. The trial produces 10 positive results and 30 negative results. Of the 10 positive results, 5 of them are false positives. Of the 30 negative results, 5 of them are false negatives.What is the specificity of this test?

a. 17%
b. 50%
c. 70%
d. 83%
e. 100%

98. A researcher wishes to compare the effects of four different antiretroviral drug combinations on the survival time of two groups of patients with AIDS; one group are IV drug abusers, the other are infants infected in utero. Each of these groups is divided into four subgroups; each subgroup is given a different drug combination. Which statistical technique would be most appropriate for analyzing the results of this study?

a. Analysis of variance (ANOVA)
b. t-test
c. F-test
d. Correlation with associated t-test
e. Chi-square

99. An investigator into the life expectancy of IV drug abusers divides a sample of patients into HIV-positive and HIV-negative groups. What type of data does this division constitute?

a. Nominal
b. Ordinal
c. Interval
d. Ratio
e. Continuous

Biostatistics SBAQ for Surgeons

Secario applies to Questions 100-105

The following questions refer to the Table below, which shows the number of cases of breast cancer occurring in a randomized clinical trial of a new drug designed to prevent the disease. In this study, 1000 healthy women between the ages of 60 and 65 were given the drug and 1000 were given the placebo for 5 years.

	DISEASE OUTCOME	
TREATMENT	Breast cancer	No breast cancer
Placebo	40	960
Drug	10	990

100. What is the absolute risk (over the 5 years duration of the study) of getting breast cancer for patients in the placebo group?

a. 4%
b. 1.6
c. 24
d. 67.5%
e. 0.4

101. What is the absolute risk (over the 5 years duration of the study) of getting breast cancer for patients in the drug group?

a. 30
b. 25%
c. 1%
d. 4
e. 0.5

102. What is the relative risk reduction of breast cancer attributable to the drug?

a. 25%
b. 50%
c. 75%
d. 100%
e. 150%

103. What is the absolute risk reduction in breast cancer attributable to the drug?

a. 1%
b. 2%
c. 3%
d. 4%
e. 5%

104. What is the absolute risk reduction in breast cancer attributable to the drug?

a. 1%
b. 2%
c. 3%
d. 4%
e. 5%

Biostatistics SBAQ for Surgeons

105. What is the number needed to treat (NNT) to prevent one case of breast cancer?

a. 2
b. 3
c. 4
d. 33.3
e. Cannot be calculated from the information given.

106. A group of patients have their albumin levels measured on hospital admission and again 3 weeks into their stay. Which one of the following statistical tests is most appropriate to test. the hypothesis that the albumin levels have changed?

a. McNemar test
b. Independent samples t-test
c. Linear regression
d. Wilcoxon's matched-pairs rank sum test
e. Paired samples t-test

107. In a randomised controlled trial of a new treatment for preventing recurrence of stroke, 1000 patients are randomised to the new treatment and 1000 to standard therapy. A total of 66 patients receiving the new treatment suffered recurrent stroke, compared to 110 in the control arm. What was the relative risk reduction?

a. 4.4%
b. 6.6%
c. 11%
d. 40%
e. 60%

108. In the design of a randomised controlled trial, to what does the term 'the power of the study' refer?

a. The size of treatment difference that the trial can be expected to detect
b. The probability of rejecting the null hypothesis that the treatments have the same effect
c. The chance that a clinically significant difference will be observed
d. The probability of a type Il error
e. The probability of establishing a statistically significant treatment effect if the true treatment difference is at a prespecified level

109. An oncologist suspects that many women diagnosed as having breast cancer also have high levels of work-related stress. What type of design would be best to test this possible association?

a. Group comparative, randomised controlled trial
b. Crossover, randomised crossover trial
c. Prospective cohort study
d. Retrospective case-control study
e. Prospective observational study

Biostatistics SBAQ for Surgeons

110. In the ETDRS study for laser treatment of diabetic macular oedema, moderate visual loss occurred in 12% of treated eyes compared to 24% of untreated eyes in 3 years.
What is the value of NNT (numbers needed to treat) in this study, approximately?

a. 2
b. 6
c. 8
d. 9
e. 12

111. A study comparing a new antibiotic to an old one for diverticulitis looks at the numbers of patients 'cured', as opposed to 'not cured' in each group using the chi-squared statistic. In the trial, the cure rates appeared to improve with the new antibiotic ($X2 = 4.2$; $P < 0.01$). What consideration should you take from this study?
a. The improvement seen with the new treatment must be clinically significant
b. The mean survival time was significantly longer for those given the new antibiotic
c. Age may have been a confounding factor
d. The trial implies that a difference in response of 4.2 times was observed
e. The results may have occurred by chance 1 time in 20

112. You are in a Pre-assessment Clinic and measure the blood pressure of patients seen that day. You note the diastolic blood pressure (DBP) in a similar group of young men. This is normally distributed with a mean of 70 mmHg and a standard deviation of 10 mmHg.
What can you conclude from this data?

a. About 50% of the men have a DBP above 70 mmHg
b. About 50% of the men have a DBP below 50mmHg
c. All the DBPs must be less than 100 mmHg
d. Should carry out a trial of much greater size
e. The distribution of DBP is skewed

113. Quantitative measures, such as blood pressure and blood test levels, have to be summarised using two descriptive statistics - a measure of central tendency and a measure of dispersion. The arithmetic mean is such a descriptive statistic.What do you understand by this term?

a. It is the average deviation
b. It is the most commonly occurring value
c. It is the nth root of the product of the observations (where n is the number of observations)
d. It is the sum of the observations divided by the number of observations
e. It is the value that divides the observations into two equal halves when they are arranged in order of increasing value

114. You are taking time out of clinical practice to undertake a research fellowship, which entails managing clinical trials. You are reading through several studies to better understand clinical trials, and come across a study that looks at the use of intravenous steroids in cardiac surgery. The trial states that it is a prospective, double-blinded, randomised control trial.Which type of bias will be reduced in this trial?

a. Attrition bias alone
b. Publication bias
c. Selection and observer bias
d. Selection bias alone
e. Selection, observer and funding bias

Biostatistics SBAQ for Surgeons

115. You are reading a research paper that has attempted to evaluate the effect of smoking on the incidence of Sinonasal polyposis. As such the researchers looked at two cohorts of matched patients, half of which had the disease, and then evaluated their smoking histories. Which one is the correct description of the level of evidence this paper represents?

a. Level I
b. Level II
c. Level III
d. Level IV
e. Level V

116. As part of his research protocol, a medical student is looking at the correlation of height to glove size in a population of surgeons. The data are assumed to be in a normal distribution pattern.Which of the following tests would be most useful?

a. Kaplan-Meier curves with log-rank analysis
b. Mann-Whitney test
c. Kruskal-Wallis test
d. Pearson's correlation co-efficient
e. Wilcoxon test

117. A group of 80 children with otitis media are randomised to either a long course or a short course of antibiotics. The number of children who had more than one attack in the following 12 months is compared. What is the most appropriate statistical test to make this comparison?

a. Analysis of variance
b. Chi-squared test
c. Student's t-test
d. Mann-Whitney U-test
e. Regression analysis

118. A group of 80 children with otitis media are randomised to either a long course or a short course of antibiotics. The number of children who had more than one attack in the following 12 months is compared. What is the most appropriate statistical test to make this comparison?

a. Analysis of variance
b. Chi-squared test
c. Student's t-test
d. Mann-Whitney U-test
e. Regression analysis

119. Medical statistics deals with the analysis, interpretation and presentation of data. A false-positive result occurs if the null hypothesis is rejected when it is actually true. What sort of error is this?

a. A experimental design error
b. Observer bias error
c. Selection bias error
d. Type I error
e. type II error

Biostatistics SBAQ for Surgeons

120. A pharmaceutical company testing a drug against postoperative pain wishes to compare the mean scores for pain between standard analgesia and the new drug. Which one of the following is the most appropriate statistical test?

a. Chi-squared test
b. Life table analysis
c. Paired t-test
d. Pearson's correlation
e. Unpaired t-test

121. A researcher plans a study to investigate whether inappropriate steroid prescription is a risk factor for bowel perforation. He decides to perform a case-control study instead of a cohort study. Which of the following best describes an advantage of a case-control study over a cohort study?

a. It provides information on a wide range of outcomes
b. It may be performed with minimal expertise
c. It is able to directly measure the incidence of a disease
d. It is able to analyse exposure to rare factors
e. The chronological sequence of events can be assessed

122. In a study of patients receiving medical treatment for thyrotoxicosis, the patient cohort receiving carbimazole (n=7000) suffered a 7% symptom recurrence rate, whereas the patient cohort receiving propylthiouracil (n=9000) suffered a 11% symptom recurrence rate, over a 5-year study period (p=0.025). The risk of symptom recurrence in an untreated population with thyrotoxicosis over this time was 25%. Which of the following percentages is the annual incidence of symptom recurrence in the treated population in this study?

a. 1.85%
b. 2.10%
c. 2.35%
d. 2.50%
e. 2.75%

123. A consultant surgeon wishes to use a non-parametric test to compare more than two groups of patients.

a. Spearman's rank analysis
b. Student's t-test
c. Mann-Whitney U test
d. Fisher's exact test
e. Kruskal-Wallis test|

124. A professor of surgery is asked by her registrar to name a commonly used statistical test that provides correlation coefficients for non-parametric comparisons.

a. Spearman's rank analysis
b. Kaplan-Meier analysis
c. Kolmogorov-Smirnov test
d. Wilcoxon signed-rank test
e. Kruskal-Wallis test

Biostatistics SBAQ for Surgeons

125. A medical student wishes to learn more about a commonly used statistical "curve' that describes survival characteristics.

a. Kaplan-Meier analysis
b. Student's t-test
c. Mann-Whitney U test
d. Spearman's rank analysis
e. Fisher's exact test

126. An academic surgeon wishes to perform a meta-analysis of studies to assess the survival outcome after endovascular aneurysm repair for ruptured abdominal aortic aneurysms. She wishes to use a specific diagram to express her survival data for this purpose.

a. Receiver operating characteristic (ROC)
b. Log-rank test
c. Kaplan-Meier curve
d. Multiple regression
e. Negative predictive value

127. A consultant ENT surgeon wishes to teach his medical students about certain cancer screening programmes in the UK. He describes a particular term that indicates the proportion of patients with normal screening test results who are truly free from disease.

a. Negative predictive value
b. Positive predictive value
c. Sensitivity
d. Specificity
e. False negative

128. An academic surgeon becomes unsure about the validity of results and conclusions from his study after he learns that the true null hypothesis had been incorrectly rejected due to statistical oversight. Which of the above terms best describes the error made in the study?

a. Type I error
b. Type II error
c. Selection bias
d. Confounding bias
e. P-hacking

129. In a trial to test the effectiveness of a new drug against the standard treatment for improving forced expiratory volume in 1 s (FEV1) among asthmatics, the members of age-matched and sex-matched pairs are randomly allocated to the new or standard treatment. The difference (new standard) in forced vital capacity (FVC) between the pairs 1 h after treatment was 0.5 ÷ 0.2 (mean ≤ standard error), $P < 0.05$. What are the chances of this difference having occurred by chance?

a. Less than one time in 100
b. Less than one time in 500
c. Less than one time in 1000
d. Less than two in 1000
e. Less than five in 100

Biostatistics SBAQ for Surgeons

130. A study is seeking to compare the weight of 88 patients at 6 months post-sleeve gastrectomy, with the weight of 56 patients at 6 months post-gastric bypass in a bariatric surgery unit 2 years ago. What is the most suitable statistical test for this comparison?

a. Unpaired t-test
b. Paired t-test
c. Friedman test
d. Analysis of variance
e. Pearson correlation coefficient

131. A 49-year-old woman presents to the General Practitioner for a preoperative assessment prior to spinal surgery. You want to perform a study comparing the heights of men and women postoperatively. What is the most appropriate statistical test to apply to the women's heights preoperatively?

a. Analysis of Variance
b. Chi squared
c. Mann-Whitney U
d. Paired T test
e. Unpaired T- test

132. In a small, randomised double-blinded trial of a new treatment in acute myocardial infarction, the mortality in the treated group was half that in the control group, but the difference was not significant. Which of the following can be concluded from this?

a. The treatment is useless
b. There is no point in continuing to develop the treatment
c. Reduction in mortality is so great that we should introduce the treatment
d. We should keep adding cases to the trial until the test for the comparison of two proportions shows a significant difference
e. We should carry out a trial of much greater size

133. The gut transit times of two different laxatives A and B in two samples of United Kingdom patients are compared using a t-test. The study is being designed as a randomised trial. Which of these statements about this study is the most likely correct?

a. It is assumed that the gut transit times in the two groups are not normally distributed
b. A confidence interval of 95% means that 95% of the sample data lie within this interval
c. The null hypothesis assumes there is no difference between the means of the two population groups
d. The number of patients must be the same in each group
e. The trial would be valid even if patients had travelled to the tropics recently

134. You are at journal club where your colleague presents a paper on the effects of a new drug on pain control. A cross-over study has been used. Which of the following is an advantage of using cross-over within patient comparison?

a. All confounders are eliminated
b. More patients are needed, which strengthens the validity of the result
c. There may be an advantage when examining the effects of long-acting drugs
d. They are always free from carry-over effects
e. They reduce errors associated with individual differences

Biostatistics SBAQ for Surgeons

135. Which of the following has the greatest impact on the positive predictive value of a test?

- **a.** Prevalence
- **b.** Subjects who are true negatives
- **c.** Specificity
- **d.** Relative risk
- **e.** None of the above

136. Which one of the following is equivalent to the pre-test probability?

- **a.** Post test odds / (1 + post-test odds)
- **b.** Pre-test odds x likelihood ratio
- **c.** The prevalence of a condition
- **d.** The incidence of a condition
- **e.** Post-test odds / likelihood ratio

137. If a sample is normally distributed which of the following is true?

- **a.** Mean = standard deviation
- **b.** Mean = standard error of the mean
- **c.** Mean = median
- **d.** Mean = variance
- **e.** The mode and standard error of the mean have the same value

138. A rapid finger-prick blood test to help diagnosis deep vein thrombosis is developed. Comparing the test to current standard techniques a study is done on 1,000 patients:

	DVT present	DVT absent
New test positive	200	100
New test negative	20	680

What is the specificity of the new test?

- **a.** 680/880
- **b.** 200/220
- **c.** 680/780
- **d.** 680/700
- **e.** 200/300

139. In a randomized study of chemotherapy drugs for hypopharyngeal cancer, a group receiving treatment A had a recurrence rate of 12.5% and a group receiving treatment B had a recurrence rate of 15%. Both groups are matched for size and length of follow up. What is the number needed to treat to prevent a recurrence?

- **a.** 2.5
- **b.** 25
- **c.** 4
- **d.** 40
- **e.** 5

Biostatistics SBAQ for Surgeons

140. Which of the statistical tests listed below is most appropriate to use with data which follows a normal distribution?

a. T Test
b. Fishers exact test
c. Spearman's rank test
d. Mann-Whitney U test
e. Bonferroni test

141. You are performing a study of weight in patients attending pre pre-operative assessment clinic for elective knee replacement. Assuming that the results are normally distributed, what percentage of values lie within two standard deviations of the mean weight?

a. 95.4%
b. 5.3%
c. 98.3%
d. 10%
e. 97.5%

142. In Intention to treat analysis, which statement is untrue?

a. It is based on the initial treatment intended.
b. Excludes those who cross over to alternative treatment arms of a trial.
c. Is different from a per protocol analysis.
d. Will affect the statistical power calculation to compare outcomes of treatment.
e. Helps to minimise observer bias.

143. In medical statistics, which of the following does a p value of 0.04 represent?

a. Risk of type 1 Error
b. Risk of type 2 Error
c. Size of power of the study
d. Sample size
e. Number of degrees of freedom

144. Considering cluster randomised trials, which of the following statements is false?

a. They consider interventions targeted at groups
b. They require increased recruitment to achieve the same level of statistical power as individual trials
c. If results are analysed on an individual basis a lower P value may be obtained
d. They are less prone to unit of analyses errors than trials involving individual observations
e. The statistical analyses for these trials is more complex than that required for trials based on individuals

145. A surgeon has conducted a piece of research and is trying to make his non-parametric data appear interesting for publication. To do this he is conducting multiple analyses of subgroup data using multiple tests. Which of the tests listed below should be considered?

a. LSD post hoc test
b. Bonferroni correction
c. Un-paired T test
d. Paired T test
e. Chi squared test

Biostatistics SBAQ for Surgeons

146. Which of the following statements relating to randomised controlled trials is false?

a. Consist of a control group recruited during the same time interval as the treatment group.
b. Are not applicable to retrospectively analysed data even if captured on a prospectively created database.
c. They require concealment of treatment throughout the duration of the study.
d. They require concealment of treatment until after randomisation.
e. They are less susceptible to researcher bias than non-randomised controlled trials.

147. A new blood test to screen patients for colon cancer is trialed on 500 patients. The test was positive in 40 of the 50 patients shown to have colorectal cancer by colonoscopy. It was also positive in 20 patients who were shown not to have colorectal cancer. What is the positive predictive value of the test?

a. 0.8
b. 0.66
c. 0.33
d. 0.1
e. Cannot be calculated

148. Which of the following tests is most appropriate for correcting against multiple statistical analyses that might provide an erroneous result (i.e. to correct against data dredging)?

a. LSD post hoc test
b. Bonferroni test
c. T Test
d. Fishers test
e. Spearman's rank test

149. Which of the following most closely describes the risk of a type I statistical error?

a. Power calculation
b. P value
c. Odds ratio
d. Relative risk
e. None of the above

150. A new bioabsorbable steroid-eluting stent, designed to maintain sinus patency, undergoes clinical trials. 100 patients are given the new stent. During a three-month period 10 of the patients have an episode of acute sinusitis. In the control group there are 300 patients who are given a placebo. In this group 50 people have an infection during the same time-period. What is the relative risk of having an infection when the new stent is used?

a. 0.8
b. 0.2
c. 1.66
d. 0.6
e. 0.06

Biostatistics SBAQ for Surgeons

151. Which of the following statements relating to quantitative data is false?

- **a.** Discrete data cannot be sub divided
- **b.** The median is less susceptible to extreme outliers than the mean
- **c.** The mean is susceptible to extreme outliers
- **d.** Data that fits the standard distribution perfectly will have a mode that is half the value of the mean
- **e.** Values obtained have a numerical scale

152. Which of the following statements relating to qualitative data is false?

- a. The data has no true numerical scale
- b. It may comprise multiple data groups
- c. May be reported using odds ratios
- d. May be reported using frequency histograms
- e. It is best analysed statistically using a students T-test when multiple factors are present

Scenario below applies to questions 151-154

A new test is developed for detecting Strep throat. In a trial 200 people are tested. The test produces 100 positive results and 100 negative results. Of the 100 positive results, 4 are false positives. Of the 100 negative results, 24 are false negatives.

153. What is the Specificity?

- **a.** 95%
- **b.** 80%
- **c.** 96%
- **d.** 76%
- **e.** 90%

154. What is the Sensitivity?

- **a.** 95%
- **b.** 80%
- **c.** 96%
- **d.** 76%
- **e.** 90%

155. What is the Negative predictive value (NPV)?

- **a.** 95%
- **b.** 80%
- **c.** 96%
- **d.** 76%
- **e.** 90%

156. What is the Positive predictive value (PPV)?

- **a.** 95%
- **b.** 80%
- **c.** 96%
- **d.** 76%
- **e.** 90%

Biostatistics SBAQ for Surgeons

157. "Blood groups" are classified as what type of data?

a. Quantitative measured data
b. Quantitative counted data
c. Ordinal data
d. Nominal data

158. "Blood pressure" is classified as what type of data?

a. Quantitative measured data
b. Quantitative counted data
c. Ordinal data
d. Nominal data

159. Staging of Laryngeal cancer, is classified as what type of data?

a. Quantitative measured data
b. Quantitative counted data
c. Ordinal data
d. Nominal data

160. What is the primary assumption of a null hypothesis?

a. There is a real difference in the underlying population within the area of investigation
b. There is no real difference in the underlying population within the area of investigation.
c. There could be a difference in the underlying population which is why the investigation is being carried out
d. There is more than likely no difference in the underlying population within the area of investigation.
e. There is more than likely a real difference in the underlying population within the area of investigation.

161. Which of the following is true in relation to a type 1 error?

a. It is also known as a beta error.
b. It is seen when we conclude that there is a difference when actually there is no real difference.
c. It is seen when we conclude that there is no difference when actually there is a difference
d. It is a false negative result.
e. It decreases as the P value increases.

162. Which one of the following is not true in relation to a type 2 error?

a. It is also known as a beta error
b. It is seen when we conclude that there is no difference when there actually is a difference
c. It Is a false-negative result.
d. It usually occurs because the sample size is too big
e. It occurs if the null hypothesis is not rejected

Biostatistics SBAQ for Surgeons

A new blood test has been developed to detect condition X. It was tested on 100 people with the following results: 45 people tested positive for X and 45 actually had the condition; 10 people tested negative for the condition but were actually found to have the condition. The remainder who tested negative were found not to have the condition.

163. What is the approximate sensitivity of the test?

a. 45%
b. 80%
c. 82%
d. 89%
e. 90%

164. What is the approximate specificity of the test?

a. 40%
b. 80%
c. 82%
d. 89%
e. 90%

165. Which of the following is not a requirement of a screening test?

a. Cost-effectiveness
b. Effective treatment exists
c. Acceptable to the population
d. Non-invasive
e. Can be done in the community

A new blood test has been developed to detect condition X. it was tested on 100 people: 50 people tested positive for X and 45 actually had the condition; 10 people tested negative for the condition but were actually found to have the condition. The remainder who tested negative were found not to have the condition.

166. What is the positive predictive value of the test?

a. 45%
b. 80%
c. 82%
d. 89%
e. 90%

167. What is the negative predictive value of the test?

a. 40%
b. 80%
c. 82%
d. 89%
e. 90%

Biostatistics SBAQ for Surgeons

A new anti-migraine medication is being developed. It was tested on 40 patients who had suffered with at least two migraines in the last 2 months which were not relieved by existing anti-migraine medications. Twenty patients received the drug and 20 patients received a placebo. Neither patients nor medical staff knew which drug was being given to which group. They were followed up over the next 2 months to see whether or not they developed migraines. Two people receiving the drug developed a migraine within this time, whilst seven patients on the placebo drug developed a migraine within this time

168. What type of study is this?

A. Single-blinded control trial
B. Double-blinded control trial
C. Open control trial
D. Cohort study
E. Experimental study

169. What is the absolute risk reduction with the new drug?

a. 10%
b. 25%
c. C, 35%
d. 65%
e. 90%

170. What is the number needed to treat?

a. 1
b. 2
c. 3
d. 4
e. 5

171. Which one of the following statements about p-value is correct?

a. It is the likelihood that a relationship under investigation is due to chance.
b. It is 100% accurate in determining a relationship.
c. It is statistically significant if less than 0.1.
d. It is statistically insignificant if less than 0.05
e. It is an example of a non-parametric test

172. An odds ratio (OR) is used as a statistical value in which of the following studies?

a. Randomised control trial
b. Cohort study
c. Case-control study
d. Ecological study
e. Experimental study

173. Which of the following statements is incorrect?

a. Point prevalence is the frequency of disease in a population at any given time
b. Period prevalence is the frequency of disease in a population over a certain time period
c. Incidence is the rate of occurrence of new disease in a population.
d. Incidence is the frequency of current disease in a population.
e. Standardised mortality ratio is the ratio of observed deaths to expected deaths in a given population.

Biostatistics SBAQ for Surgeons

174. Which of the following statements is incorrect?

a. A cohort study identifies a group of people and follows them through time to see if a certain exposure causes a certain condition. This is then compared with a similar group of people not exposed to the exposure under investigation.
b. A case-control study compares a group of people with a certain condition with a group of people without that condition and looks back through time to evaluate any potential causative factors
c. In an ecological study, the smallest unit of observation is a population or community. Ecological fallacy is a recognised error in this sort of study.
d. Vaccination is an example of primary prevention.
e. Screening is an example of tertiary prevention.

175. Which of the following statements is incorrect?

a. In a single-blinded control trial, members of either the treating team or the patient group know which medication they are receiving.
b. Material published in a foreign language and excluded in a meta-analysis is an example of publication bias.
c. Confounding factors are factors outside of the search criteria that may affect results.
d. In a study with a 95% confidence interval of 1.32-1.49, $P<0.05$ is statistically insignificant.
e. Randomisation is a method of reducing selection bias

176. A case-control study has been completed comparing 3000 female breast cancer sufferers in the UK who have also has hip fractures. The study (see table below) states that more patients will have fractures whilst receiving oestrogen receptor antagonist (ORA).

	Received tamoxifen	Never prescribed ORA
Neck of femur fracture	516	296
No neck of femur fractures	984	1204
Total number of patients	1500	1500

177. What is the observed odds ratio?

a. 0.45
b. 0.62
c. 1
d. 2
e. 2.13

178. A randomised control trial has just been published comparing the effect of a new antihypertensive medication, ZXA®, on Afro-Caribbeans. In the study the results have shown that patients who were given ZXA had better controlled blood pressure. Which of the following statements best describes the hypothesis testing?

a. The null hypothesis is true (Ho).
b. One could always apply a one-tailed test to this study.
c. The null hypothesis has to be rejected.
d. The alternative hypothesis is true (H,).
e. The alternative hypothesis is true.

Biostatistics SBAQ for Surgeons

179. A recent study multi-centred, randomised control trial with a total recruitment of 25,000 has been published comparing the effects of a new bowel cancer drug. It states that there is a 45% reduction in 5-year mortality from bowel cancer in patients who have taken the new drug. The P-value was set before the study at P< 0.05. The study is asking you to reject the null hypothesis. You reject the null hypothesis, but a review article is published a month later which states that a significant number of patients were actually lost to follow-up and that there was in fact no difference in mortality between the two groups.

What type of error have you committed?

a. A type 1 error, as the alpha level not calculated
b. A type 1 error
c. A type 2 error
d. A type 1 and 2 error
e. No error at all

180. You are asked to by your consultant to see if you can apply some statistics to a study he has been conducting, which is comparing exam results of module 2 following implementation of new teaching aid (e-modules. He wants to compare these results with the exam performance of the same students in module 1. After analysing the data and establishing it is normally distributed, which one of the following would you apply?
a. Paired t-test
b. Wilcoxon signed-rank test
c. Unpaired t-test
d. Mann-Whitney U-test
e. One-way analysis of variance (ANOVA)

181. A workforce analysis survey is being conducted at a major computer company. They are categorising the workforce into gender, age and ethnicity. The line manager is trying to decide which statistical test is best to apply. Which one of the following should he choose?

a. One-way ANOVA
b. Chi-squared test
c. Paired t-test
d. Unpaired t-test
e. Wilcoxon signed rank test

182. In a small double-blind study of pain following minor surgery, patients were randomly allocated to receive either an analgesic or a placebo 1 hour preoperatively. The patients were then asked to rate their pain on a scale of 1 to 4, 1 being no pain and 4 being severe pain. Which of the following is the most appropriate statistical test for analysing the results?

a. Mann-Whitney U-test
b. Unpaired Student's t-test
c. Fisher's exact test
d. Chi-squared test
e. One-way analysis of variance

Biostatistics SBAQ for Surgeons

183. Research has been carried out into a new drug, measuring the effect on heart rate before and after administration of the drug in a population of healthy volunteers. Which of the following statistical tests would be most appropriate?

a. Analysis of variance (ANOVA)
b. Wilcoxon signed rank
c. Friedman
d. Kruskal-Wallis
e. Fisher's exact test

184. Which of the following terms best describes the precision to which a collection of samples represents the population from which they are drawn?

a. Standard error of the mean (SEM)
b. Standard deviation
c. Variance
d. Mean
e. Central tendency

185. A new blood test is developed to identify patients who develop ventilator-associated pneumonia. The test is described by the company representative as having a high positive Predictive value (PPV). Which of the following statements is most likely to be accurate regarding the value of this test?

a. The area under the receiver operating curve (AUROC) is likely to be close to 0.5.
b. The negative predictive value (NPV) will probably also be high.
c. The PPV will be similar when tested in two different populations.
d. The PPV can be calculated by dividing the number of true positives by the number of false-positive results
b. The PPV can be calculated by dividing the number of true positives by the sum of all positive results.

186. A range in which the true value lies
What is the single best term for this definition?
a. Confidence interval
b. Forest plot
c. Median
d. Mode
e. Sensitivity

187. To determine the difference between the actual and expected number of reactions
What is the single best term for this definition?
a. Chi-square
b. Fisher's test
c. Mann-Whitney U test
d. P value
e. Student's t-test

188. Ability of the test to pick up disease
What is the single best term for this definition?
a. Negative predictive value
b. Positive predictive value
c. P value
d. Sensitivity
e. Specificity

Biostatistics SBAQ for Surgeons

189. One hundred patients with heavy postmenopausal bleeding had a pipelle endometrial biopsy taken prior to a hysterectomy to test the ability of the biopsy to detect endometrial cancer.

		Endometrial Cancer	
		Affected	Non-Affected
Biopsy	**Positive**	20	10
	Negative	30	40

What is the Specificity of the endometrial biopsy?

a. 20%
b. 40%
c. 50%
d. 80%
e. 100%

190. A new antenatal screening test to assess the level of cervical effacement in threatened preterm delivery has been developed. In trials this test has a high degree of specificity. What is the definition of the term?

a. The proportion of false positives that is correctly identified by the test
b. The proportion of people with negative results who are correctly diagnosed by the test
c. The proportion of people with positive results who are correctly diagnosed by the test
d. The proportion of true negatives that is correctly identified by the test
e. The proportion of true positives that is correctly identified by the test

191. The table below shows the results of a new test for the detection of a disease

		Disease	
		Affected	Non-Affected
Test	**Positive**	a	b
	Negative	c	d

How would you calculate the sensitivity of the new test?

a. a/a+d
b. a/a+c
c. b/b+c
d. b/b+d
e. d/a+b+c

192. In evidence-based medicine, what level of evidence is provided by meta-analyses of randomised controlled trials?

a. 1a
b. 1b
c. 2a
d. 2b
e. 3

Biostatistics SBAQ for Surgeons

193. A study is conducted to compare the efficacy of 5% hypertonic saline and 20% mannitol in reducing intracranial pressure (ICP) in patients with traumatic brain injury. Mean ICP values are available for two groups before and after the intervention. Assuming that the ICP values are normally distributed (Gaussian distribution), which of the following statistical tests would be most appropriate in comparing the two groups?

a. Wilcoxon signed rank test.
b. Student's t-test.
c. ANOVA paired test.
d. Paired student's t-test.
e. Mann-Whitney U test.

194. A paper has been published looking at the effectiveness of bystander cardiopulmonary resuscitation (CPR) in out of hospital cardiac arrest. The paper pooled together data from six different trials in order to achieve a sample size of over 3,000 patients over a 10-year period. That data was then analysed collectively. Which single type of study design best describes this type of trial?

a. Case-control study
b. Meta-analysis
c. Prospective cohort study
d. Randomized control trial
e. Retrospective cohort study

195. The following results were obtained from a randomized control trial for a new diagnostic test, called DJW, for diagnosis of subarachnoid hemorrhage (SAH). Which is the single most appropriate description of the trial results?

	SAH	
	Present	**Absent**
DJW—Positive	8	10
DJW—Negative	32	30

SAH, subarachnoid haemorrhage.

a. DJW has a negative predictive value (NPV) of 75%
b. DJW has a positive predictive value (PPV) of 50%
c. DJW has a sensitivity of 20%
d. DJW has a sensitivity of 75%
e. DJW has a specificity of 20%

196. A psychiatrist devised a short screening test for depression. An independent blind comparison was made with a gold standard for diagnosis of depression among 200 psychiatric outpatients. Among the 50 outpatients found to be depressed according to the gold standard, 35 patients were positive for the test. Among 150 patients found not to be depressed according to the gold standard, 30 patients were found to be positive for the test.
Corrected to the nearest decimal place:

a. The sensitivity was 80%
b. The specificity was 70%
c. The positive predictive value was 70%
d. The negative predictive value was 88.9%
e. The prevalence of depression was 80%

Biostatistics SBAQ for Surgeons

A total of 360 patients participated in a randomised controlled trial designed to compare the effectiveness of drug X in reducing deaths with a placebo. Out of 120 patients in the treatment group, 12 patients died within three years. Out of 240 patients in the control group, 48 patients died within three years.

197. The relative risk reduction(of treatment compared with control group) is

a. 10%
b. 20%
c. 25 %
d. 50%
e. 67%

198. The number needed to treat (NNT) to prevent 1 death is

a. 10
b. 20
c. 1000
d. 100
e. 15

199. In a randomised controlled trial, which of the following is reduced by randomisation?

a. Ascertainment bias
b. Selection bias
c. Recall bias
d. Publication bias
e. Bias in handling dropouts

Biostatistics SBAQ for Surgeons

A recent study compared the diagnostic performance of T2-weighted MRI only vs. T2W and diffusion-weighted MRI together for the assessment of residual tumour in patients who underwent chemotherapy for oesophageal cancer.

	Tumour	No Tumour
+ive Test	37	6
-ive Test	2	6

200. What is the Sensitivity?

a. 95%
b. 90%
c. 80%
d. 50%
e. 86%

201. What is the specificity?

a. 95%
b. 90%
c. 80%
d. 50%
e. 86%

202. What is the positive predictive value?

a. 95%
b. 90%
c. 80%
d. 50%
e. 86%

203. Blood pressure measurements are made in a group of children with pituitary hormone disorders and in age- and sex-matched control pairs. The study aims to investigate whether pituitary hormone disorders are associated with altered blood pressure.

a. Two-sample t-test
b. Paired t-test
c. Mann-Whitney U-test
d. Kruskal-Wallis test
e. Chi-square test

Biostatistics SBAQ for Surgeons

204. Developmental tests are applied to determine whether children who were admitted to intensive care in the neonatal period are more likely to have delayed development at age 5 years than those who were not.

a. Two-sample t-test
b. Paired t-test
c. Mann-Whitney U-test
d. Kruskal-Wallis test
e. Chi-square

205. Blood pressures (assumed normally distributed) are compared between 5-year-olds from four different, clearly defined racial backgrounds.

a. Two-sample t-test
b. Paired t-test
c. Mann-Whitney U-test
d. One-way analysis of variance
e. Kruskal-Wallis analysis of variance
f. Regression analysis
g. Chi-square

206.

Biostatistics SBAQ for Surgeons

Answers